USAARL Report No. 98-10

The Impact of Task Length on Multi-Attribute Task Battery (MATB) Performance During Sleep Deprivation

By

John A. Caldwell, Jr.
Stephanie Ramspott
Susan J. Gardner

Aircrew Health and Performance Division

January 1998

U.S. Army Aeromedical Research Laboratory
Fort Rucker, Alabama 36362-0577

Notice

Qualified requesters

Qualified requesters may obtain copies from the Defense Technical Information Center (DTIC), Cameron Station, Alexandria, Virginia 22314. Orders will be expedited if placed through the librarian or other person designated to request documents from DTIC.

Change of address

Organizations receiving reports from the U.S. Army Aeromedical Research Laboratory on automatic mailing lists should confirm correct address when corresponding about laboratory reports.

Disposition

Destroy this document when it is no longer needed. Do not return it to the originator.

Disclaimer

The views, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other official documentation. Citation of trade names in this report does not constitute an official Department of the Army endorsement or approval of the use of such commercial items.

Human use

Human subjects participated in these studies after giving their free and informed voluntary consent. Investigators adhered to AR 70-25 and USAMRMC Reg 70-25 on Use of Volunteers in Research.

Reviewed:

MORRIS R. LATTIMORE, JR.
Colonel, MS
Director, Aircrew Health &
Performance Division

Released for publication:

For JOHN A. CALDWELL, Ph.D.
Chairman, Scientific Review
Committee

CHERRY L. GAFFNEY
Colonel, MC, SFS
Commanding

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release, distribution unlimited
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) USAARL Report No. 98-10	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION U.S. Army Aeromedical Research Laboratory 6b. OFFICE SYMBOL (If MCMR-UAD	7a. NAME OF MONITORING ORGANIZATION U.S. Army Medical Research and Materiel Command
6c. ADDRESS (City, State, and ZIP Code) P.O. Box 620577 Fort Rucker, AL 36362-0577	7b. ADDRESS (City, State, and ZIP Code) Fort Detrick Frederick, MD 21702-5012
8a. NAME OF FUNDING / SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
62787A	30162787A879	OC	175

11. TITLE (Include Security Classification)
(U) The Impact of Task Length on Multi-Attribute Task Battery (MATB) Performance During Sleep Deprivation

12. PERSONAL AUTHOR(S)
John A. Caldwell, Stephanie Ramspott, and Susan J. Gardner

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month,	15. PAGE COUNT
Final	FROM TO	1998 January	20

16. SUPPLEMENTAL NOTATION
USAARL Report 98-10 has been accepted for publication in Behavioral Research Methods, Instruments, & Computers.

17. COSATI CODES

FIELD	GROUP	SUB-GROUP
06	10	
05	07	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Task duration, sleep deprivation, performance, MATB

19. ABSTRACT (Continue on reverse if necessary and identify by block number)
Several factors affect the sensitivity to sleep deprivation when taking computerized tests. Two of these are task complexity and duration. Complexity can improve sensitivity by increasing the demands required to complete the task, but it can also decrease sensitivity by improving the subject's motivation. The effects of task duration are more predictable in that longer tests generally are more susceptible to depravation effects than shorter ones. The impact of test duration on an interesting, but complex aviation simulation was examined here. By breaking down data from 30-minute MATB administrations into the first, second, and third 10-minutes of performance, it was shown that tests shorter than 30 minutes underestimate the impact of sleep loss on performance. This was especially evident in measures of time-out and tracking errors.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Chief, Science Support Center	22b. TELEPHONE (Include Area Code) (334) 255-6907 22c. OFFICE SYMBOL MCMR-UAX-SI

DD Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE

Unclassified

Table of contents

Page

Background . . . 1

Objectives . . . 2

Methods . . . 3
- Subjects . . . 3
- Apparatus . . . 3
- Procedure . . . 3
- Testing Schedule . . . 5

Data Analysis . . . 5

Results . . . 6
- Fuel Management . . . 6
- Communications . . . 6
- Systems Monitoring . . . 7
- Tracking . . . 13

Discussion . . . 14

References . . . 15

List of figures

1. The layout of the MATB as it was presented on the computer screen . . . 5
2. Effects of interval and time on time-out errors in communications . . . 7
3. Effects of testing time (with intervals collapsed) on time-out errors in communications and both reaction time measures and time-out errors in systems monitoring . . . 8
4. Effects of interval and time on the reaction time for lights in systems monitoring . . . 9
5. Effects of interval and time on the standard deviation of reaction times for lights in systems monitoring . . . 10
6. Effects of interval and time on time-out errors for dials in systems monitoring . . . 11
7. Effects of interval and time on time-out errors for lights in systems monitoring . . . 11
8. Effects of interval and time on RMS errors in tracking . . . 13

List of tables

Effects of task duration (first, second, and third 10-minute interval) on performance . . . 12

Background

It is well accepted that sleep deprivation produces a wide array of performance decrements including increased reaction times, decreased vigilance and alertness, distortions in cognition and perception, and degraded mood (Krueger, 1989). However, a review of individual studies in the literature shows substantial variation in the effects of sleep loss from one investigation to another. Discrepancies among studies are likely due to a number of design and/or methodological factors rather than to differences in the effects of sleep loss per se.

One factor which contributes to disparate findings across sleep-deprivation studies (which is of particular importance in partial deprivation or napping studies) is the number of minutes between the time of awakening and the time of testing. Wilkinson and Stretton (1971) for instance, found that reaction time, arithmetic calculations, and psychomotor coordination were particularly impaired when subjects were tested within 15 minutes of being awakened from sleep. This is a phenomenon referred to as sleep inertia, or the tendency toward postsleep impairment and disorientation that is especially severe after a brief period of sleep inserted into a prolonged period of wakefulness (Dinges, 1989). If testing is always performed immediately upon awakening from sleep, it could erroneously be concluded that, from a performance standpoint, even a full-night's sleep is worse than total sleep deprivation. Recent investigations in our Laboratory attest to this problem by showing that strategic napping (compared to total sleep deprivation) substantially impaired mood states immediately upon awakening, despite the fact that napping markedly improved mood, alertness, and performance later in the deprivation period (Caldwell et al., 1997).

A second factor is the amount of feedback given to sleep-deprived subjects. Generally, feedback seems to improve performance by augmenting motivation levels. Wilkinson (1961) found that knowledge of results is especially important for maintaining the performance of sleep-deprived subjects. Feedback significantly attenuated the effects of 1 night of total sleep deprivation on performance of a five-choice test of serial reaction.

A third factor is the type of task chosen for the research approach. It has been found that some tasks are more likely than others to be affected by sleep deprivation. Wilkinson (1964), for instance, reported that while serial reaction time and vigilance tests were significantly degraded by sleep loss, rote learning and various games were not. He attributed these differences to such factors as task complexity and the subjects' interest in the tests. Generally, the more complex and less interesting the task, the more performance suffered after sleep loss. However, if the task was both complex *and* interesting, the subjects' interest levels apparently provided enough intrinsic incentive to overcome the degrading effects of complexity--a finding which has received support from Elsmore et al. (1995), and others.

A final factor which accounts for discrepancies across sleep deprivation studies is test duration. It has been found that performance on most tasks can be adequately maintained for

brief periods of time (i.e., 5-10 minutes) even in subjects who are significantly sleep deprived, but when task durations are extended to 20-30 minutes, performance deteriorates rapidly (Wilkinson, 1969a; Wilkinson, Edwards, and Haines, 1966). Recently completed research, in which the Synthetic Work Battery (SYNWORK) was administered to sleep-deprived aviators in a sustained work scenario (Caldwell et al., 1994; Caldwell et al., 1995; Caldwell et al., 1997), tends to support this contention. In the first two studies, the effects of placebo versus dextroamphetamine on the performance of sleep-deprived subjects were examined using 10-minute administrations of SYNWORK placed at 4-hour intervals. Although this task is one which requires substantial sharing of cognitive resources (among simultaneously presented memory, arithmetic, visual monitoring, and auditory monitoring components), no drug-related performance differences were found in either of the 2 studies, with the exception of a single drug-by-session interaction on 1 variable out of 11 (performance on only 1 subtask out of the 4 was affected). In a later study of the effects of strategic naps versus total sleep deprivation, the duration of SYNWORK was increased to 20 minutes and the same basic testing schedule was repeated. This time, results indicated intervention-related effects on three of the four subtasks despite the fact that only basic composite scores (instead of the more sensitive individual measures used in the earlier studies) were analyzed. Thus, lengthening the test duration from 10 minutes to 20 minutes appeared to substantially increase task sensitivity.

In summary, differences in elapsed times from awakening, knowledge of results, task sensitivity, and task duration, as well as other factors, may account for inconsistencies in the published literature regarding the impact of sleep deprivation on both mood and performance. It is unlikely that sleep loss is such a minor stressor that it produces decrements only in specific subgroups or in certain circumstances. Instead, it is more plausible that sleep deprivation exerts a universally negative effect on humans which is difficult to reliably quantify because of differences across studies. In fact, a meta-analysis by Pilcher and Huffcutt (1996) suggests that methodological inconsistencies result in a general underestimation of the impact of sleep deprivation in humans. The precise effects of test-specific factors remain unclear because complex interactions no doubt exist. For instance, while it is fairly certain that increasing task duration will improve task sensitivity to sleep deprivation, it is less obvious whether this will apply equally to both complex and simple tasks.

Objectives

The present study was conducted to more clearly delineate the effects of task duration on sensitivity to sleep deprivation in circumstances where a complex, multi-task test battery, specifically, the Multi-Attribute Task Battery (MATB) is used. To address the "time-on-task" issue, 30-minute blocks of test administrations were used. These blocks were subsequently broken into three successive 10-minute intervals for analysis purposes. The factor of sleep inertia (time since awakening) was controlled by testing subjects continuously, with no sleep during the latter part of a 38-hour period of sustained wakefulness (subjects all had been awake

for at least 13 hours prior to the first test). To minimize confounds regarding knowledge of results, feedback was not provided to any of the participants (a future study will establish the impact of this factor in a similar context). The MATB was selected based on the fact that 1) it requires a high level of cognitive resource sharing, and 2) it has been rated by subjects tested in this Laboratory over the past several years as a good, face-valid method for assessing aviator performance. Both rated aviators and flight students find the MATB challenging and interesting since it is a complex aviation simulation.

Methods

Subjects

Eighteen males (rated aviators and flight students) between the ages of 22 and 31 (mean age of 24.4) participated after pre-screening and medical evaluation. One subject's data were excluded due to computer problems. Candidate subjects who used tobacco; currently took medications other than ibuprophen, aspirin, acetominophen, or sodium naproxin; consumed more than three 8-ounce cups of coffee or five 12-ounce caffeinated soft drinks per day; or experienced a current, medical disorder (including sleep abnormalities) were excluded. Females were not tested because none volunteered. Subjects refrained from alcohol or other drugs during the protocol.

Apparatus

The MATB is a computer-based, aviation-related, synthetic task battery which was initially developed by NASA researchers (Comstock and Arnegard, 1992). The test was implemented on a 486 desktop computer equipped with a game card (Gamecard 3, C.H. Products), a voice synthesizer card (Soundblaster 16, Creative Lab.) with stereo headphones (Sony), a joystick (Advance Gravis Computer Tech. LTD), and a standard keyboard and color monitor.

Procedure

The MATB included a resource (fuel) management task, a communications task, a systems monitoring task, and an unstable tracking task, each of which was presented in a separate quadrant of the computer screen (see figure 1). Subjects were instructed they would be completing tasks designed to assess their performance in a simulated flight environment. Specifically, they were told they would be performing a tracking task while simultaneously monitoring system status and communication channels and managing fuel resources. Subjects were not provided with any instructions about the relative importance of any one task over another task, nor were they advised about how they should attempt to divide their attention among the different subtasks. Instructions for the individual subtasks were as follows:

The system monitoring task will require attending to the four gauges marked F1, F2, F3, and F4 and the two boxes marked F5 (usually green) and F6 (usually blank) on the computer screen. Use the corresponding keys to manipulate the boxes and gauges. Press the *F5* key immediately if the F5 box is no longer green. The F6 box should always be blank. Press the *F6* key if the F6 box turns red. The pointers in the gauges need to be within one tick mark above or below the mid-line. Press the corresponding keys immediately if the pointers move beyond the one-tick-mark range. When this out-of-range status is correctly identified, the pointer will move immediately back to the mid-line and a bar at the bottom of the gauge will be illuminated in yellow.

The tracking task should be executed using the joystick to keep a target in the center of its window within the dotted lines that form a rectangle. This is an attempt to simulate the demands of manual control.

The communications task simulates receiving audio messages from Air Traffic Control through a set of headphones. You will respond only to the call sign "NGT504" and make appropriate frequency changes on a Navigation and/or Communication radio. Your call sign will be displayed at the top of the Communications window and you will need to discriminate your call sign from other three-letter, three-number combinations. A command to change frequency will only be repeated once. Use *Up* and *Down* arrow keys on the keyboard to move from "NAV1" through "COM2". Use *Left* and *Right* arrow keys on the keyboard to change frequency. The *Left* arrow key will decrease the frequency; the *Right* arrow key will increase the frequency. Every keystroke will result in a 0.2 MHZ change in radio frequency. Press the *Enter* key to acknowledge the completed frequency adjustments.

The goal of the resource (fuel) management task is to maintain tanks A and B at 2500 units each, which is indicated by numbers below the tanks. This desired level is also indicated by a tick mark in the shaded bar on the sides of the two tanks. The shaded region surrounding the tick mark represents acceptable performance. The resource management task can be accomplished by turning on or off any of the pumps labeled 1 through 8. Periodically, a pump failure will occur and the pump will turn red; this pump cannot be turned back on until the red light goes out. The process of transferring fuel is accomplished by activating the pumps using corresponding number keys. Pressing the number key a second time will turn that particular pump off. When the pump is actively transferring fuel, it will turn green, and the direction of transfer is indicated by arrow keys. The maximum capacity for tanks A and B is 4000 units, and for tanks C and D, the maximum capacity is 2000 units. The remaining two supply tanks have an unlimited capacity. The flow rates for each pump are shown in the "Pump Status" window. However, you do not manipulate the pump status; it is simply used as a gauge.

In the resource (fuel) management task, either pump 2 or pump 4 failed once every 2 minutes. In the systems-monitoring task, there was either a dial or light indication requiring a response from the subject 3 times per minute. In the communications task, radio messages were delivered at a rate of 2 messages per minute. A response was required for half of these messages.

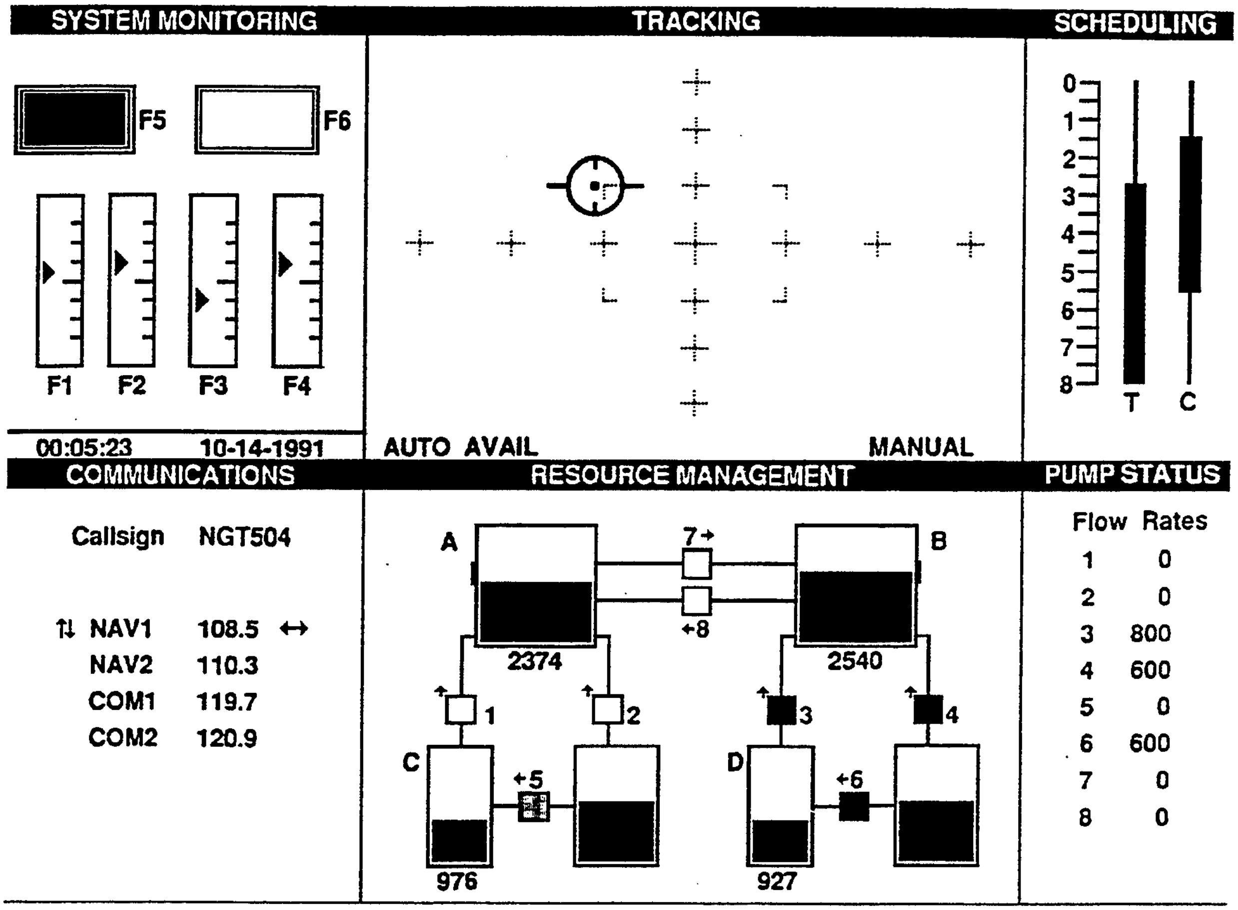

Figure 1. The layout of the MATB as it was presented on the computer screen.

Testing Schedule

All test subjects resided in the Laboratory throughout the experiment. After initial training, there was a full 10-hour night of sleep prior to actual testing. Upon awakening (at 0800), subjects completed predeprivation testing at 4-hour intervals (at 0910, 1310, and 1710) prior to an evening in which they engaged in light physical activity, watched television, and interacted with staff members, but were not allowed to sleep. At 0110, subjects began deprivation testing, and other sessions followed at 0510, 0910, 1310, and 1710. Each test administration was 30 minutes long. A variety of test activities (mood and alertness tests) occurred between each of the MATB administrations, and a complete schedule has been published elsewhere (Caldwell et al., 1997).

<u>Data Analysis</u>

A computer program was used to extract a number of performance indices from each MATB administration. The deviation of fuel levels from the ideal value of 2500 units in the fuel-

management task; the reaction time, standard deviation of reaction time, and time-out errors in the communications task; the reaction times to lights and dials, the standard deviations of reaction times to lights and dials, and time-out errors for lights and dials in the systems-monitoring task; and the root mean square (RMS) tracking errors in the unstable tracking task were saved for analysis.

The data from each iteration of the test were segmented into three intervals (first 10 minutes, second 10 minutes, and third 10 minutes) for each of the four tasks (fuel management, systems monitoring, etc.). Two-way analyses of variance (ANOVA) were used to examine differences attributable to interval (first, second, and third 10 minutes) and time (predeprivation--0910, 1310, and 1710; deprivation--0110, 0510, 0910, 1310, and 1710). Both predeprivation and deprivation data were included in each analysis in order to ensure that any observed changes in task sensitivity were a function primarily of sleep deprivation rather than simply time of day (Chmiel, Totterdell, and Folkard, 1995).

Results

Fuel Management

There were no significant main effects or interactions on the mean deviation of fuel units from 2500 (the ideal) in the fuel management task. Apparently, the within-subjects variability in responding was so great it overshadowed any deprivation effects.

Communications

There was an interaction between interval and time (test session) on time-out errors in the communications task ($F(14,224)=2.60$, $p=.0016$), but neither the reaction time for correct responses nor the standard deviation of reaction times was similarly affected. The interval-by-time interaction was due to variations in the pattern of differences among the various testing times within the first, second, and third 10-minute intervals ($p<.05$). Pairwise contrasts indicated none of the predeprivation conditions differed from one another at any of the intervals, but instead, the differences all occurred between predeprivation (PD) and sleep-deprivation (SD) sessions or within the deprivation sessions themselves (see figure 2). In the first 10 minutes, time-out errors were fewer at PD 0910 and PD 1710 than at SD 1310 or SD 1710, and fewer at PD 0910 than at SD 0510. In the second 10 minutes, time-out errors were fewer at PD 0910, PD 1310, and PD 1710 than at SD 0910; errors were fewer at PD 0910 than at SD 1310; and errors were larger at both the SD 0910 and SD 1310 times than at SD 1710. In the third 10 minutes, time-out errors were lower at all three PD sessions than at SD 0910 and SD 1310; and, in addition, they were lower at PD1710 than at SD 0110 and SD 0510, whereas they were higher at both SD 0910 and SD 1310 than at SD 1710. Generally, sleep-deprivation effects became more apparent as a function of test duration. A time main effect on time-out errors ($F(7,112)=4.08$,

p=.0005) showed that errors (with intervals collapsed) were less numerous at both PD 0910 and PD 1310 than at SD 0910; errors were less numerous at PD 1310 than at SD 1310; and less at PD 1710 than at any of the sleep deprivation sessions, with the exception of SD 1710 ($p<.05$). This main effect, along with the time main effects from other subtests, is shown in figure 3. There were no main effects on the interval factor.

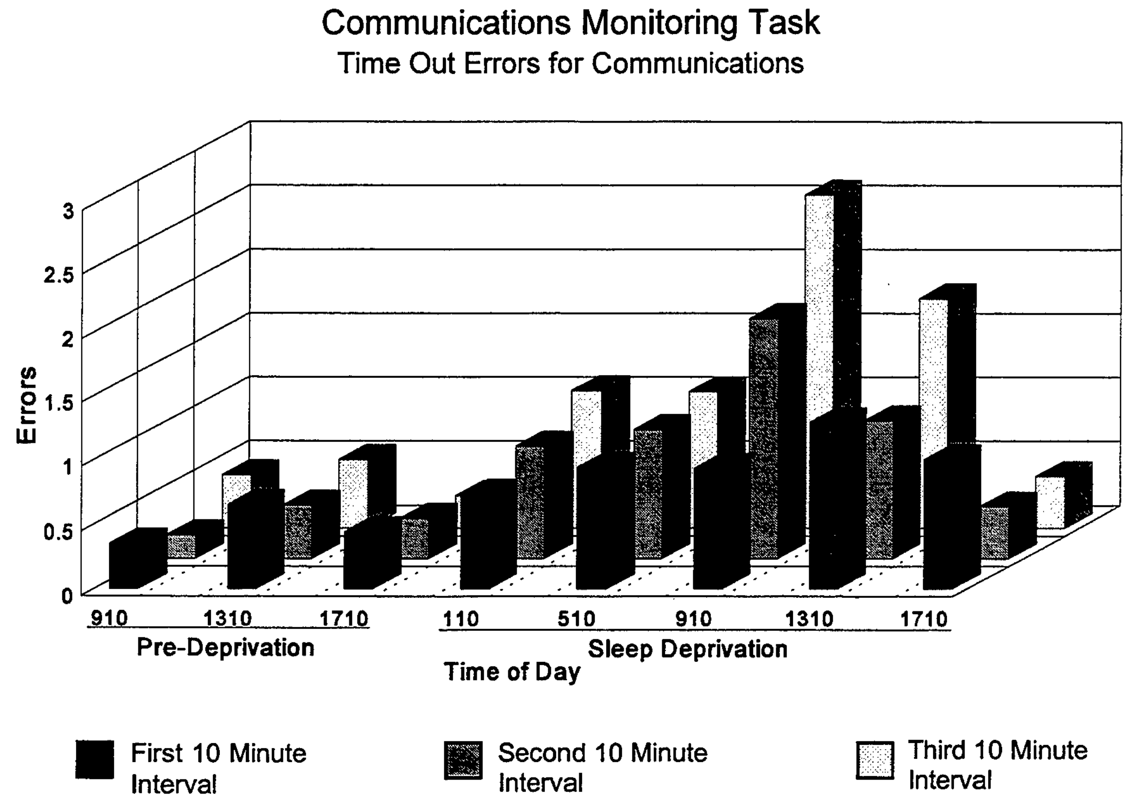

Figure 2. Effects of interval and time on time-out errors in communications.

Systems Monitoring

There were interval-by-time interactions on the reaction time for lights ($F(14,224)=2.10$, $p=.0126$), the standard deviation of reaction times for lights ($F(14,224)=2.57$, $p=.0019$), time-out errors for lights ($F(14,224)=1.80$, $p=.0391$), and time-out errors for dials ($F(14,224)=2.64$, $p=.0014$). Analysis of simple effects indicated there were differences within each interval (first, second, and third 10 minutes) across the predeprivation and deprivation sessions for both of the reaction time measures ($p<.05$). Generally, deprivation-related changes became more noticeable as the duration of the task increased. There were no differences among the test sessions which preceded sleep deprivation. Within the first 10-minute interval, although none of the PD

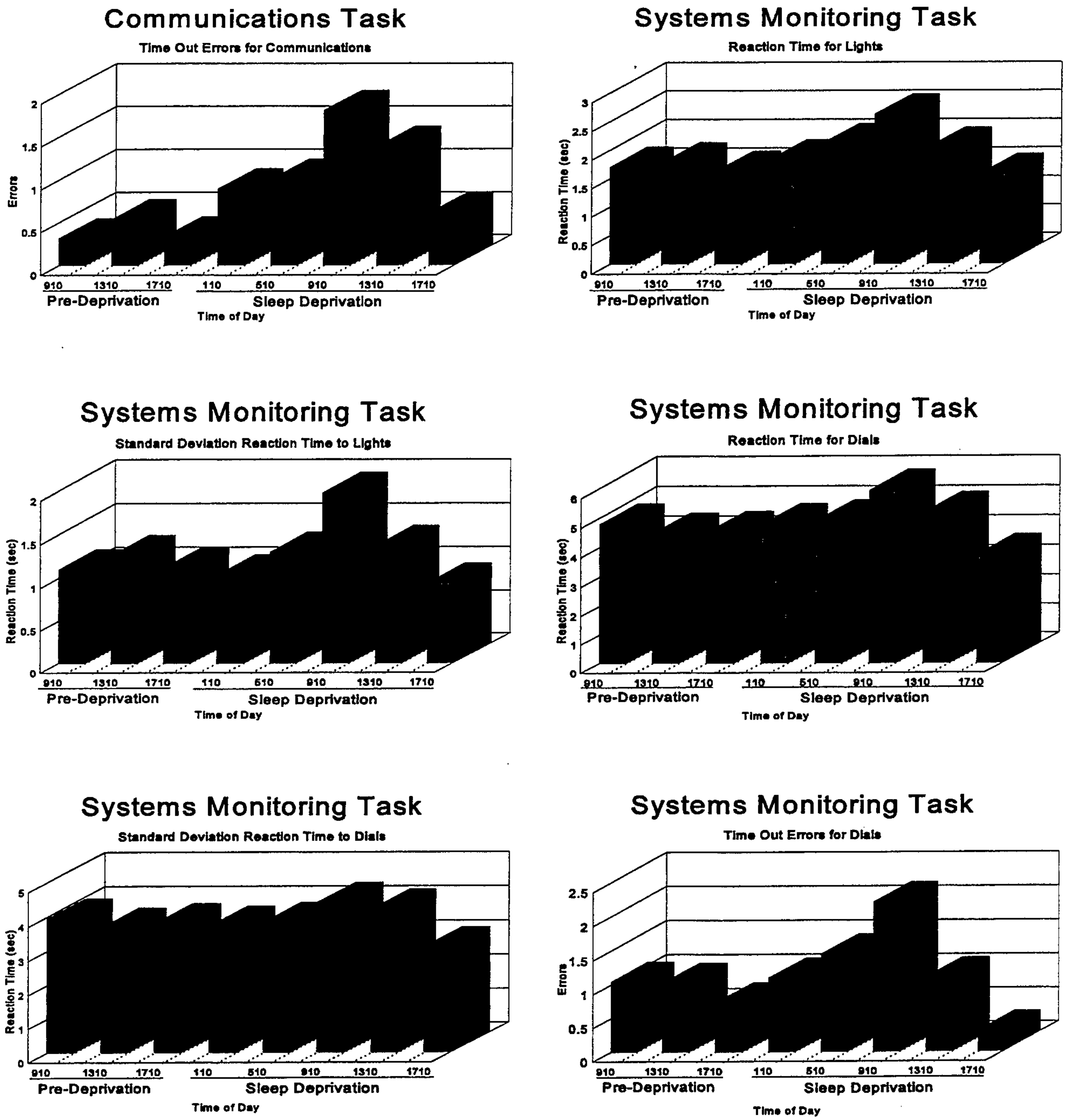

Figure 3. Effects of testing time (with intervals collapsed) on time-out errors in communications and both reaction time measures and time-out errors in systems monitoring.

sessions differed from one another, reaction times were longer at SD 0910 than at PD 0910 or PD 1710, whereas they were shorter at SD 1710 than at PD 0910 or PD 1310. Within the SD sessions, reaction times were slower at SD 0510 and SD 0910 than at SD 1310 or SD 1710. Also, they were slower at SD 0110 than at SD 1710 ($p<.05$). Within the second 10-minute

interval, reaction times were slower at SD 0910 than at any of the PD sessions or any of the SD sessions, with the exception of SD 0510. Reaction times at SD 0510 were longer than those at PD 1710 or SD 1710; and reaction times at SD 1310 were longer than those at PD 1710, SD 0110, or SD 1710 ($p<.05$). Within the third 10-minute interval, reaction times were slower at SD 0910 than at any of the PD sessions or any of the SD sessions, with the exception of SD 1310; slower at SD 0510 and SD 1310 than at PD 0910 or PD 1710; and slower at SD 1310 than at SD 1710 ($p<.05$). These reaction-time effects are shown in figure 4. The pattern of the standard-deviation data was somewhat similar, but the overall number of effects was smaller. Within the first 10-minute interval, almost all of the other sessions (with the exception of SD 0510 and SD 1310) were found to be more variable than the SD 1710 session, while all of the SD sessions were less variable than the SD 0910 session ($p<.05$). Within the second 10-minute interval, the SD 0510, SD 0910 and SD 1310 sessions were more variable than the SD 1710 session; and all of the PD sessions, as well as the SD 0110 and SD 1710 sessions, were less variable than the SD 0910 session. SD 1310 was more variable than PD 1710 or SD 0110. Within the third 10-minute interval, the variability in SD 1710 had increased so that only SD 0910 was greater, but the variability at SD 0910 was higher than the variability at any of the PD or SD sessions. Also, the variability in reaction times at SD 0510 was greater than the variability at PD 0910, PD 1710, or SD 0110 ($p<.05$). These effects on the standard deviation of reaction times are depicted in figure 5.

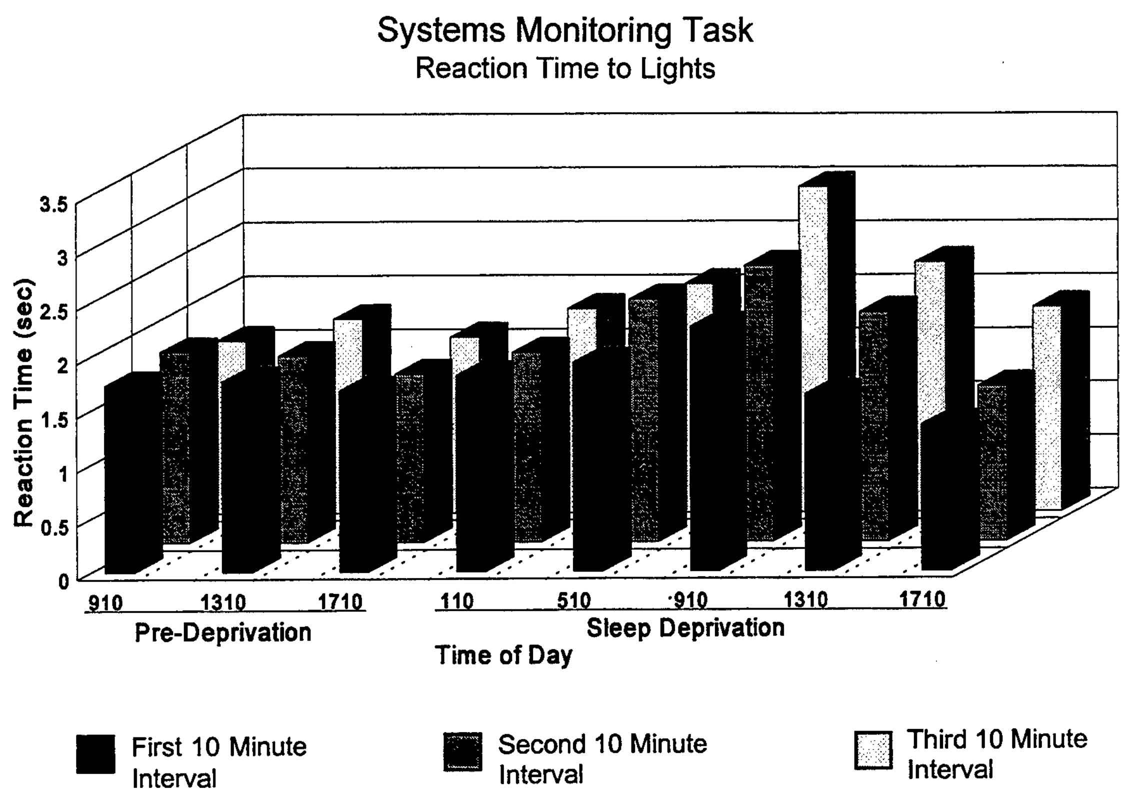

Figure 4. Effects of interval and time on the reaction time for lights in systems monitoring.

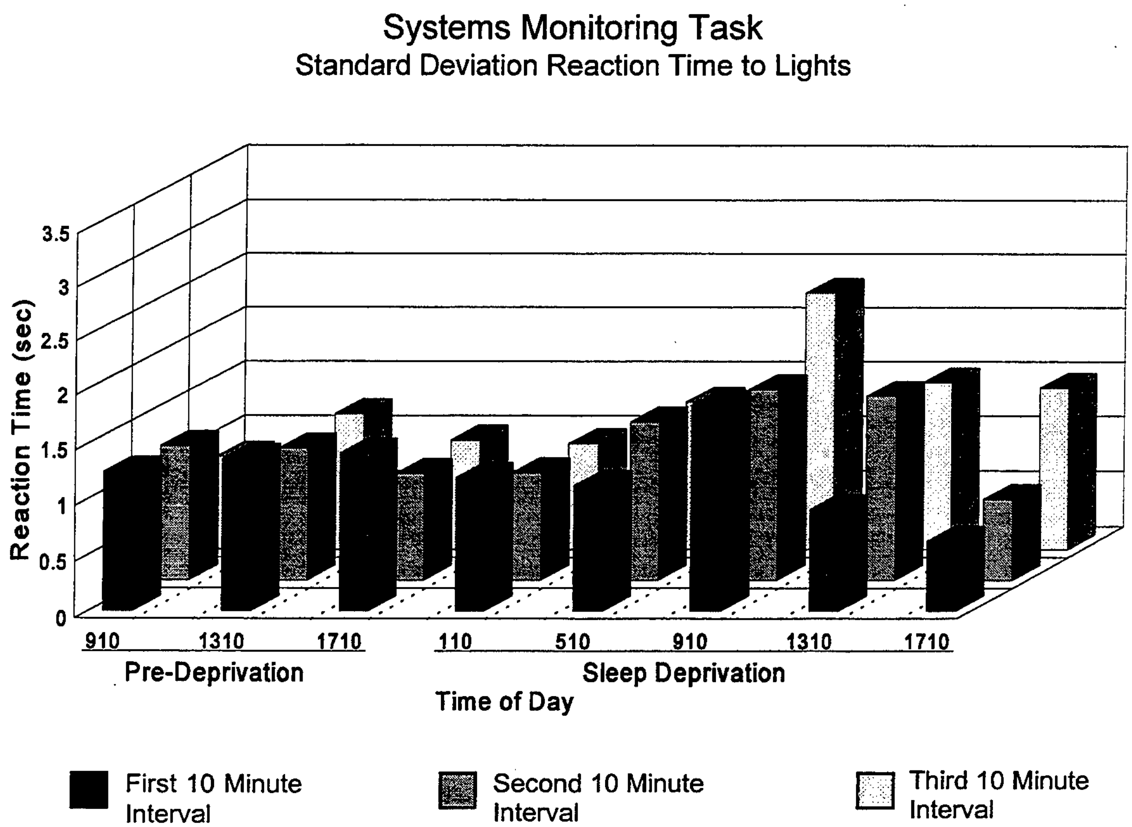

Figure 5. Effects of interval and time on the standard deviation of reaction times for lights in systems monitoring.

The interval-by-time interaction for time-out errors for dials was attributable to the fact that there were differences among the testing times only within the first 10-minute interval and the third 10-minute interval ($p<.05$), but not the second. As was the case with the reaction time measures, the general trend was toward larger discrepancies between predeprivation and deprivation sessions as task duration increased from 10 to 30 minutes. There were no differences among the predeprivation performance scores. Within the first 10-minute interval, the only effects were due to smaller time-out errors during the SD 1710 session than during all of the other sessions ($p<.05$), with the exception of PD 1710 and SD 1310 (where no differences occurred). Within the second 10-minute interval, there were no significant effects as has already been noted. Within the third 10-minute interval, the general trend toward poorer performance at 0910 during sleep deprivation (SD 0910) became evident as errors at this time were significantly more frequent than errors at any of the PD sessions or any of the SD sessions, with the exception of SD 1310 ($p<.05$). The effects on time-out errors for dials are shown in figure 6. The interaction on time-out errors for lights was attributable to the presence of a marginally significant difference across the testing times only in the third 10-minute interval ($p=.0551$), but not the first or second. As can be seen in figure 7, this effect was primarily due to inordinately poor performance at SD 0910 (most clearly evident only after 30 minutes of task duration).

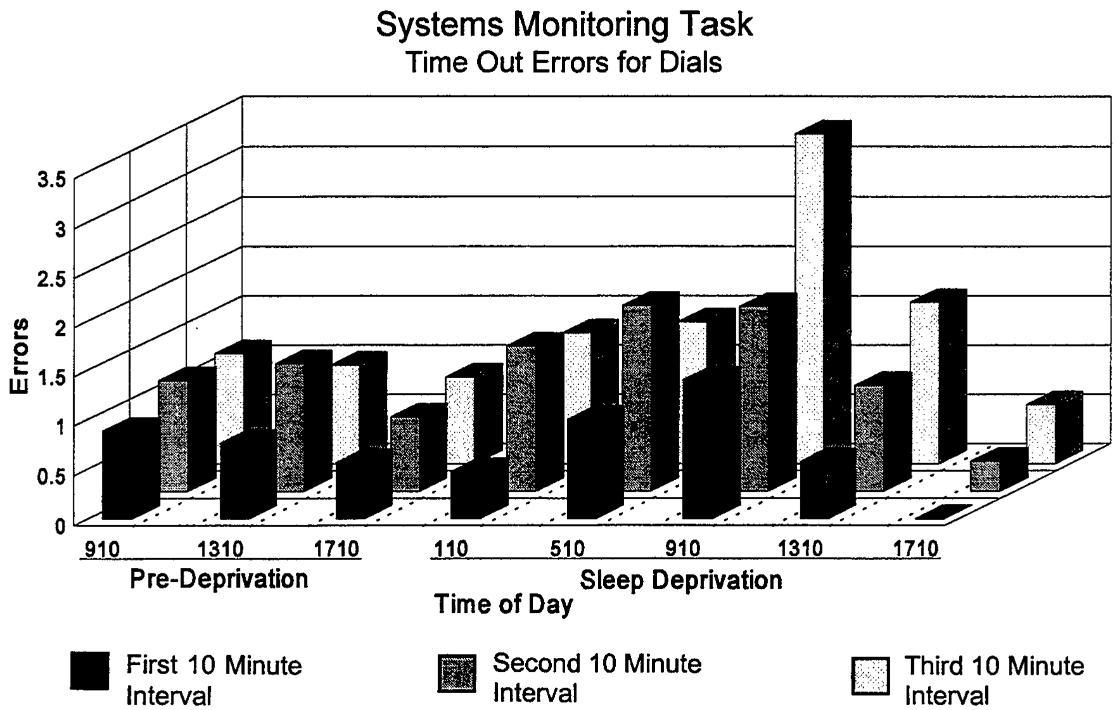

Figure 6. Effects of interval and time on time-out errors for dials in systems monitoring.

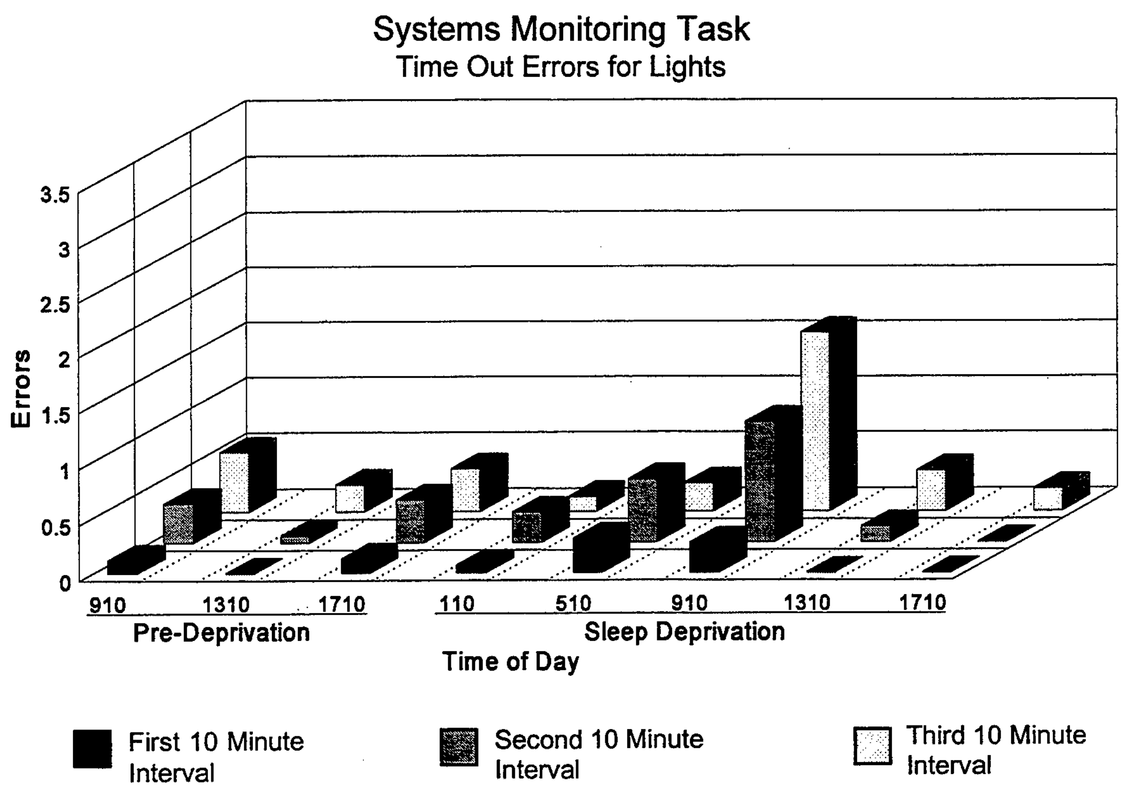

Figure 7. Effects of interval and time on time-out errors for lights in systems monitoring.

In addition to these interactions, there were time main effects on the reaction time to lights ($F(7,112)=7.36$, $p<.0001$) and dials ($F(7,112)=6.32$, $p<.0001$), the standard deviation for correct responses to lights ($F(7,112)=5.76$, $p<.0001$) and dials ($F(7,112)=3.53$, $p=.0018$), and the time-out errors to dials ($F(7,112)=2.80$, $p=.0100$). The reaction times to both lights and dials were slower at SD 0910 than at any of the predeprivation or deprivation sessions. In addition, reaction time to lights was slower at SD 0510 than at PD 1710, and slower at SD 1310 than at either PD 0910, PD 1710, or SD 1710. Reaction time to dials was slower at SD 0510 than at SD 0110 or SD 1710, and faster at SD 1710 than PD 0910, SD 0110, or SD 1310. The reaction time to dials also was slower at the 1310 deprivation session than at the predeprivation 1310 session ($p<.05$). The standard deviation for correct responses to lights was larger at SD 0910 than at any other session (predeprivation or deprivation); larger at SD 0510 than at SD 0110; and larger at SD 1310 than at SD 1710 ($p<.05$). The main effect on standard deviation for correct responses to dials was due to the small variability at SD 1710 in comparison to all of the other sessions (predeprivation and deprivation) and a slight reduction in variability from the first to the second predeprivation sessions ($p<.05$). The main effect on time-out errors for dials was due to more frequent errors at SD 0910 than at PD 1710 or SD 1710 while the errors at SD 1710 were lower than those at PD 0910 and SD 1310 ($p<.05$).

There were main effects on the interval factor for the reaction time to lights ($F(2,32)=6.25$, $p=.0051$) and dials ($F(2,32)=8.94$, $p=.0008$), and the time-out errors for lights ($F(2,32)=4.63$, $p=.0171$) and dials ($F(2,32)=10.86$, $p<.0003$). In every case, performance was worse (i.e., longer reaction times and increased errors) in the third 10-minute interval than in the first; in every case except for the reaction time for lights, performance was worse in the second 10-minute interval than in the first; and for reaction time to dials, performance was worse in the third interval than in the second ($p<.05$). These interval effects are shown in the table below.

Table.

Effects of task duration (first, second, and third 10-minute interval) on performance.

Variate	First 10 minutes	Second 10 minutes	Third 10 minutes
Reaction time lights	1.78	1.90	2.04
Reaction time dials	4.47	4.76	5.08
Time outs lights	0.11	0.35	0.46
Time outs dials	0.71	1.22	1.42

Tracking

There was an interval-by-time interaction ($F(14,224)=3.62$, $p<.0001$), a time main effect ($F(7,112)=6.71$, $p<.0001$), and an interval main effect ($F(2,32)=19.74$, $p<.0001$) on tracking RMS errors. Although there were differences among the test sessions at every interval length, the interaction was attributable to a deprivation-related worsening in the pattern of performance as a function of time on task ($p<.05$). There were no differences in tracking performance within the predeprivation period. In the first 10-minute interval, performance in the SD 0910 session was worse than all of the other sessions, with the exception of SD 0510; and the SD 0510 session was worse than PD 1310 and SD 0110, as well. Performance at SD 1310 was poorer than performance at PD 1310 or SD 1710 ($p<.05$). In the second 10-minute interval, performance at SD 0910 was worse than all of the other sessions, without exception. Although performance at SD 0510 was more impaired as compared to the first 10-minute interval, tracking skill at SD 1310 was much lower than at any of the other sessions, with the exception of SD 0510 and SD 0910 ($p<.05$). In the third 10-minute interval, tracking at SD 0910 was poorer than tracking at every session, with the exception of SD 1310 (at which time tracking had deteriorated almost as much as it had at SD 0910), and tracking at SD 0510 was worse than tracking at PD 1310 and SD 0110. In addition, performance was lower at SD 1310 than at PD 0910, PD 1310, PD 1710, SD 0110, and SD 1710 ($p<.05$). These effects are shown in figure 8.

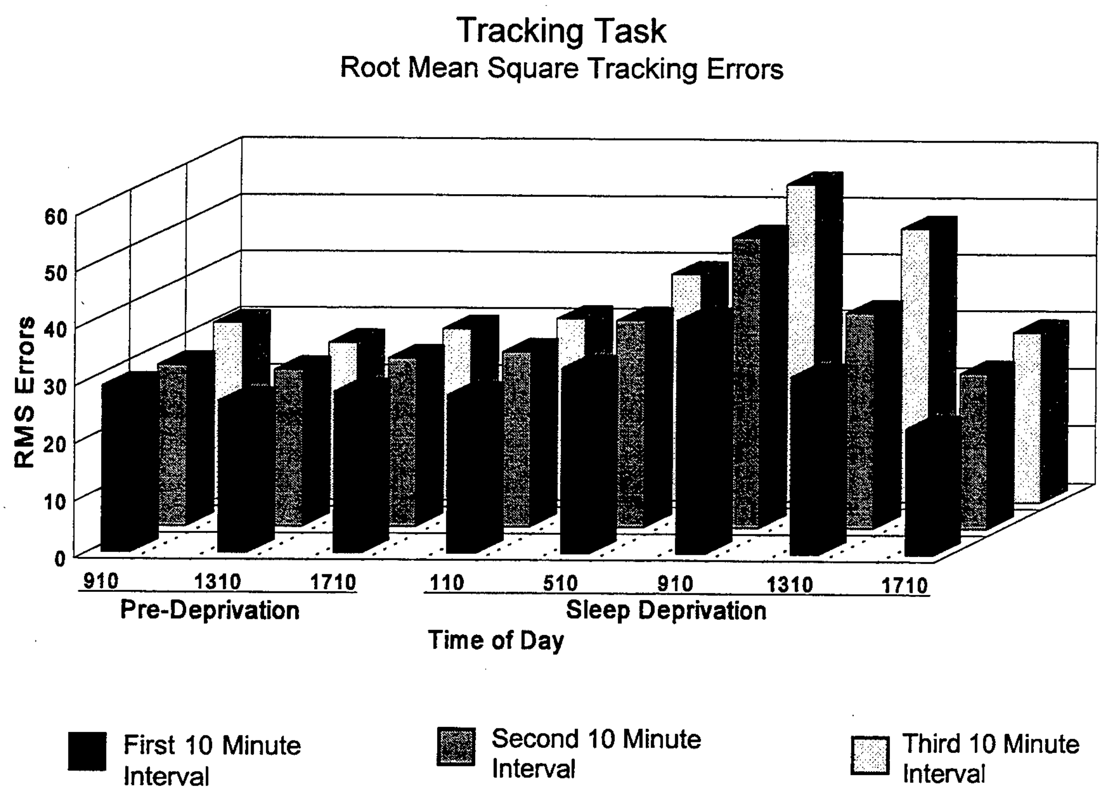

Figure 8. Effects of interval and time on RMS errors in tracking.

The main effect on the time factor occurred because tracking was worse at SD 0910 than at any other time and worse at SD 1310 than at all three PD sessions, SD 0110, and SD 1710. In addition, performance at SD 0510 was worse than at PD 1310 or SD 0110 ($p<.05$). The main effect on the interval factor was due to a progressive worsening of performance from the first 10-minute interval to the third 10-minute interval (all comparisons were statistically significant).

Discussion

This investigation showed that performance on a 30-minute aviation simulation task during 38 hours of continuous wakefulness declined most severely in the mid morning after 1 night of sleep loss. Furthermore, the study revealed that if a shorter task duration had been selected, the overall impact of sleep deprivation on performance would have been seriously underestimated (or missed altogether), despite the fact that a rather complex sharing of mental resources was required by the selected task (the MATB). An examination of 11 performance variates (i.e., reaction-time measures, errors, etc.) indicated that task durations of 10, 20, and 30 minutes clearly were differentially sensitive to the effects of sleep loss. Particularly noticeable were the increases in reaction times (responding to communications calls and warning lights), time-out errors (responding to warning lights and dial deviations), and tracking errors, all of which became more pronounced as task duration increased from 10 minutes to 30 minutes.

Such findings are consistent with Wilkinson's (1969a) report that even subjects who are significantly sleep deprived are quite capable of short periods (10-15 minutes) of mental concentration that far exceed their capacity for normal work shifts (which may last 8 or more hours). Thus, a reliance on very short laboratory tests to predict performance decrements in an actual work setting can be dangerous, especially in contexts where the tolerance for error is quite small (such as in aviation). It may seem possible to compensate for brief test periods by implementing more complex tasks--a possibility that was partially examined here. However, it is evident that such a strategy can backfire and result in decreased test sensitivity because complex tasks tend to be more interesting than boring tests and, therefore, less susceptible to fatigue-induced decrements (Elsmore et al., 1995; Wilkinson, 1969b). This is particularly true in sleep-deprivation paradigms where sleepy subjects briefly can be returned to "normal" if sufficiently aroused. However, as Wilkinson (1969b) has stated, "the subject's problem is that he habituates to arousing stimuli [quickly] and so this normality cannot be maintained for very long, and hence [the] time factor beats him" (p. 34).

The findings from the present study reaffirm the importance of using increased task duration to heighten test sensitivity, especially when attempting to gain an understanding of the effects of a stressor such as sleep deprivation. It is concluded that even when fairly demanding, face valid simulation tasks are employed, at least 30 minutes of continuous performance will be required before generalizations to "real-world" work performance can even begin to be made.

References

Caldwell, J. A., Jones, R. W., Caldwell, J. L., Colon, J. A., Pegues, A., Iverson, L., Roberts, K., Ramspott, S., Sprenger, W. D., and Gardner, S. J. 1997. The efficacy of hypnotic-induced prophylactic naps for the maintenance of alertness and performance in sustained operations. Fort Rucker, AL: United States Army Aeromedical Research Laboratory. USAARL Technical Report No. 97-10.

Caldwell, J. A., Caldwell, J. L., Crowley, J. S., Darling, S. R., Wallace, S., Woodrum, L. C., and Jones, H. D. 1995. Sustaining female UH-60 helicopter pilot performance with Dexedrine in sustained operations: A simulator study. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Technical Report No. 95-27.

Caldwell, J. A., Caldwell, J. L., Crowley, J. S., Jones, H. D., Darling, S. R., Wallace, S. L., Woodrum, L. C., and Colon, J. A. 1994. Effects of dextroamphetamine on helicopter pilot performance: A UH-60 simulator study. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. USAARL Technical Report No. 94-43.

Chmiel, N., Totterdell, P., and Folkard, S. 1995. On adaptive control, sleep loss, and fatigue. Applied Cognitive Psychology. 9: S39-S53.

Comstock, J. R., and Arnegard, R. J. 1992. The multi-attribute task battery for human operator workload and strategic behavior research. Hampton, NY: Langley Research Center, National Aeronautics and Space Administration. NASA Technical Memorandum No. 104174.

Dinges, D. F. 1989. Napping patterns and effects in human adults. Sleep and alertness: Chronobiological, behavioral, and medical aspects of napping. D. F. Dinges and R. J. Broughton (eds). New York: Raven Press. 171-204.

Elsmore, T. F., Hegge, F. W., Naitoh, P., Kelly, T., Schlangen, K., and Gomez, S. 1995. A comparison of the effects of sleep deprivation on synthetic work performance and a conventional performance assessment battery. San Diego, CA: Naval Health Research Center. NHRC Report No. 95-6.

Krueger, G. P. 1989. Sustained work, fatigue, sleep loss, and performance: A review of the issues. Work & Stress. 3: 129-141.

Pilcher, J. J., and Huffcutt, A. I. 1996. Effects of sleep deprivation on performance: A meta-analysis. Sleep. 19: 318-326.

Wilkinson, R. T. 1961. Interaction of lack of sleep with knowledge of results, repeated testing, and individual differences. Journal of Experimental Psychology. 62: 263-271.

Wilkinson, R. T. 1964. Effects of up to 60 hours' sleep deprivation on different types of work. Ergonomics. 17: 175-186.

Wilkinson, R. T. 1969a. Some factors influencing the effect of environmental stressors upon performance. Psychological Bulletin. 72: 260-272.

Wilkinson, R. T. 1969b. Sleep deprivation: Performance tests for partial and selective sleep deprivation. Progress in Clinical Psychology, Vol. 3, Dreams and Dreaming. L. Edwin and B. F. Riess (eds). New York: Grune & Stratton. 34.

Wilkinson, R. T., Edwards, R. S., and Haines, E. 1966. Performance following a night of reduced sleep. Psychonomic Science. 5: 471-472.

Wilkinson, R. T., and Stretton, M. 1971. Performance after awakening at different times of night. Psychonomic Society. 23: 284-285.